CONFERENCE ON QUALITY BASED INFORMATION SYSTEMS

Preliminary Program

Place: Pocantico Hills, New York
Time: May 22 and 23, 1978

DAY 1

9:00 a.m. - 9:15 a.m.	Introduction - J. KNOWLES, President, Rockefeller Foundation, New York City, New York
9:15 a.m. - 10:00 a.m.	Scientific Communication and the Scientific Literature - J. ZIMAN, F.R.S., Professor of Physics, Bristol University, Bristol, England
10:00 a.m. - 10:45 a.m.	Quality Based Information Systems: An Ecological Approach - W. GOFFMAN, Professor, School of Library Science, Case Western Reserve University, Cleveland, O.
10:45 a.m. - 11:00 a.m.	COFFEE
11:00 a.m. - 11:45 a.m.	Need for Quality Based Information Systems in Industry, Government and Academia - V.A. NEWILL, Director, Environmental Health, Exxon Corporation, Linden, N.J.
11:45 a.m. - 12:30 p.m.	The National Library of Medicine and Quality Information Systems - M. CUMMINGS, Director, National Library of Medicine, Bethesda, Md.
12:30 p.m. - 2:00 p.m.	LUNCH
2:00 p.m. - 2:45 p.m.	The Journal as a Quality Filter - A. RELMAN, Editor, New England Journal of Medicine, Boston, Mass.
2:45 p.m. - 3:30 p.m.	The Science Citation Index as a Quality Filter - E. GARFIELD, Chairman, Institute for Scientific Information, Philadelphia, Pa.
3:30 p.m. - 4:15 p.m.	Analysis of a Medical Literature: A Case Study - K.S. WARREN, Director, Health Sciences, Rockefeller Foundation, New York City, N.Y.
Evening	DINNER Speaker: Daniel J. Boorstin, Librarian of Congress

DAY 2

9:30 a.m. - 10:15 a.m.	Use of Literature Analysis in Establishing Funding Policy - E. CONNELL, Assoc. Director, Health Sciences, Rockefeller Foundation, New York City, N.Y.
10:15 a.m. - 11:00 a.m.	Small Quality Based Information Systems and Libraries, Particularly for Developing Countries - T. SARACEVIC, Professor, School of Library Science, Case Western Reserve University, Cleveland, O.
12:00 noon - 1:30 p.m.	LUNCH
1:30 p.m. - 2:00 p.m.	SUM UP
2:00 p.m. - 3:00 p.m.	General Discussion Outlining Next Steps

The Exponential Proliferation of the Scientific Literatures

The already vast and rapidly increasing quantity of scientific information has been overwhelming to many scientists. This has resulted in widespread disaffection with the complex scientific information system that has evolved over the past three centuries. The extent of this discontent has been manifested by the suggestions of authorities such as J. D. Bernal and T. Fox (editor of the Lancet) to drastically alter the system. The most extreme approaches have involved elimination of scientific journals and their replacement with centralized computerized information systems. In contrast to this attitude is the statement of J. M. Ziman that "The invention of a mechanism for the systematic publication of fragments of scientific work may well have been the key event in the history of modern science." He further suggested that this device has been the secret of the success of western science since the 17th century.

An approach to encompassing the vast output of scientific information has been the development of secondary literature citation sources, the most notable in the biomedical field being the computerized MEDLARS system established by the National Library of Medicine. Until recently technological limitations required MEDLARS to be maintained at the level of entering and cross-indexing the citations from 2200 of the approximately 7000 biomedical journals now being published; in a recent 1.5 year period this resulted in 527,000 articles being entered into the computer banks. The present computers can be replaced by machines having the capacity to enter far more citations, cross-index them more extensively and even include abstracts of papers. However, it is clear where this approach will lead.

The Ecology of Scientific Communication

Information scientists and bibliometricians are only beginning to attempt to understand the complex interrelationships involved in the generation and transmission of information. Mathematical models have been derived from the field of epidemiology on the basis of the belief that the transmission of information is similar to the spread of infectious diseases. The biological relationship most closely resembling that of the formal system of scientific communication is a process involving an intermediate host, in this case being the scientific journal. Thus, the author produces a manuscript which if accepted by a journal proliferates and is transmitted to many other potential authors(Figure 1). This ecological model is clearly a cyclic one and it has been suggested that any drastic change in one phase of the cycle might destroy the entire system.

Good, Neuter, and Bad Literature

One approach to managing the "literature explosion" without altering the ecology is a quality-based information system. This is based on the concept that there are good, neuter and bad papers. Good papers may be defined as those that significantly advance knowledge. They can be judged on the basis of how the investigations are performed (easily determined in the traditional materials and methods sections of scientific papers) and their place as a piece of the edifice of verified information in a given field. These papers comprise no more than 10 percent of all those written, and probably considerably less. For instance, E. Garfield has observed that 70 percent of biomedical papers are never cited even once. The neuter

literature is in the vast middle area, which probably comprises more than 80 percent of published papers. These publications serve an ephemeral function of transmitting information (i.e. a review in a state or hospital medical journal) or are poorly executed studies which are essentially harmless. The bad literature constitutes no more than 10 percent of the total and may be defined as information that is wrong or misleading, either because of poor execution or interpretation. An important factor in determining whether a paper is actively bad is the proliferation of good literature necessary to contradict it. The recent spate of literature on Vitamin C is an example of this situation.

Qualitative Approach to the Scientific Literature

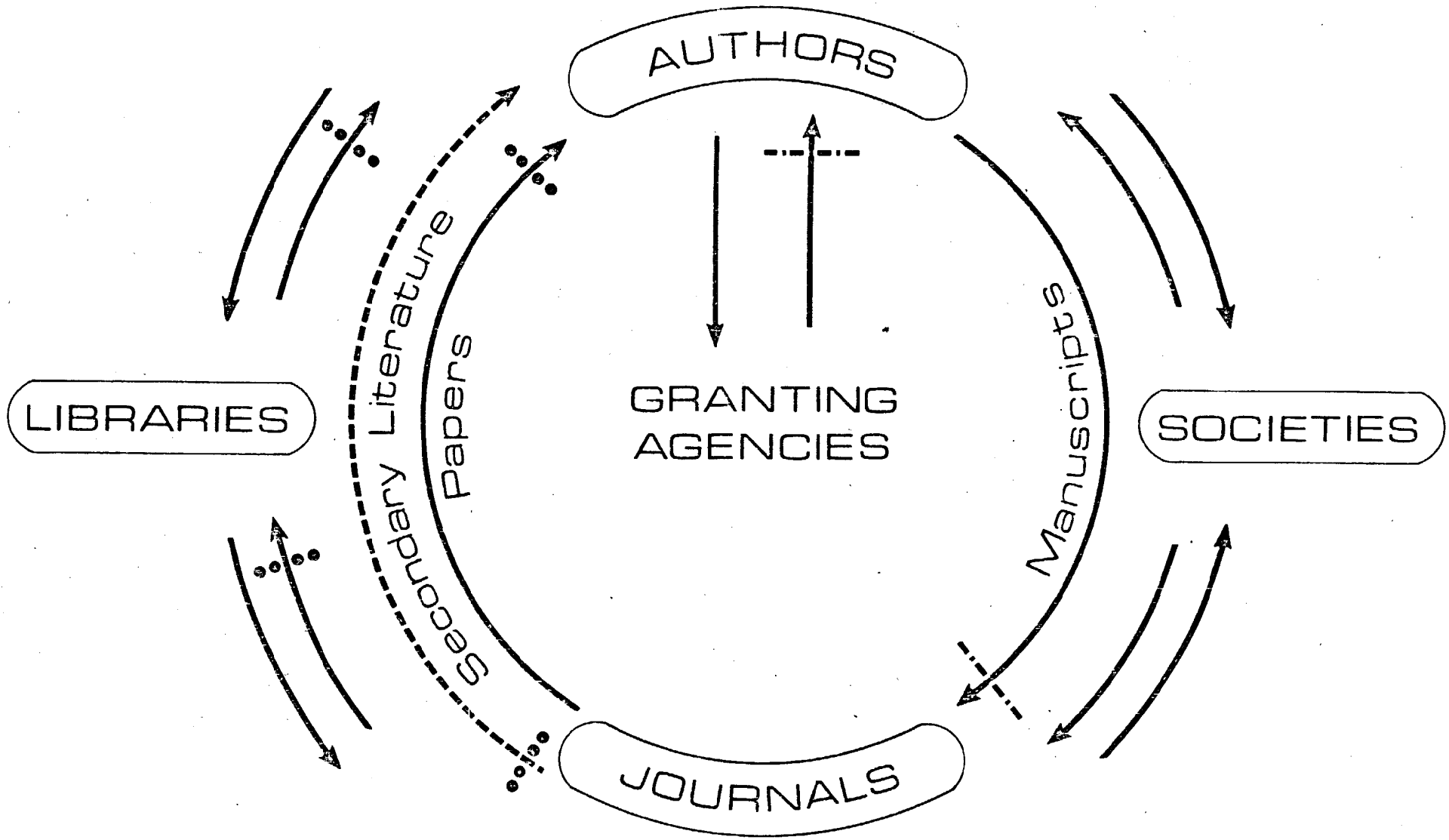
If the above assumptions are reasonably correct, then it can immediately be seen that identification of the literature of quality can drastically reduce the amount of information impinging upon the scientist and markedly improve its usefulness. It should be clearly stated at this point that no qualitative information system can even approach perfection, i.e., identifying 100 percent of the good literature. It is possible, however, to envision systems that would encompass 90 percent of the information of significant value. Each individual can then approach the total information base in various ways to try to locate the rest of the good material, recognizing that an exponential effort is involved in adding each 1 percent of good new information. Such a quality approach would not tamper with the ecology of the information cycle (Figure 1). Furthermore, quality filtration systems might have a positive feedback resulting in the quantitative decrease in the poorest segments of the system. While

the most obvious way of identifying quality would be by a large jury of scientific peers, in general this would be impractical. Thus other means must be found.

The Rockefeller Foundation Involvement

A research program focusing on quality-based information systems is being planned. In addition, a two day conference entitled Quality Based Information Systems will be held at the Rockefeller Archives Center in May, 1978. The participants will include scientists, scholars, editors, librarians and foundation officers. As a consequence of this conference new areas of investigation should be revealed especially in the area of quality-based information systems.

Figure 1



-·-·- Present quality filters

•••• Proposed quality filters